

OIPE

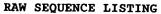
RAW SEQUENCE LISTING DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

Input Set : A:\FimH Peptides.ST25.txt
Output Set: N:\CRF3\07232001\1900575.raw

```
ENTERED
      3 <110> APPLICANT: Langermann, Solomon
              Revel, Andrew
      5
              Auguste, Christine
      6
              Burlein, Jeanne
      8 <120> TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
     10 <130> FILE REFERENCE: 469201-549
C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/900,575
C--> 12 <141> CURRENT FILING DATE: 2001-07-06
     12 <150> PRIOR APPLICATION NUMBER: US/60/216,750
     13 <151> PRIOR FILING DATE: 2000-07-07
     15 <160> NUMBER OF SEQ ID NOS: 64
     17 <170> SOFTWARE: PatentIn version 3.0
     19 <210> SEQ ID NO: 1
     20 <211> LENGTH: 837
     21 <212> TYPE: DNA
     22 <213> ORGANISM: E. coli
     24 <400> SEQUENCE: 1
     25 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatgtt
                                                                               60
     27 tatgtaaacc ttgcgcccgt cgtgaatgtg gggcaaaacc tggtcgtgga tctttcgacg
                                                                              120
     29 caaatetttt gecataacga ttateeggaa accattacag actatgteac actgeaacga
                                                                              180
     31 ggctcggctt atggcggcgt gttatctaat ttttccggga tcgtaaaata tagtggcagt
                                                                              240
     33 agetateett teectaeeac cagegaaaeg eegegegttg tttataatte gagaaeggat
                                                                              300
     35 aageegtgge eggtggeget ttatttgaeg eetgtgagea gtgegggggg agtggegatt
                                                                              360
     37 aaagcagget cattaattge egtgettatt ttgegacaga ccaacaacta taacagegat
                                                                              420
     39 ggtttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc
                                                                              480
    41 tgcgatgctt ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgccg
                                                                              540
    43 attectetta cegittatig tgcgaaaage caaaacetgg ggtattacet eteeggeaca
                                                                              600
    45 accgcaggtg cgggcaactc gattttcacc aataccgcgt cgttttcacc cgcgcagggc
                                                                              660
     47 gtcggcgtac agttggcgcg caacggtacg gttattccag cgaataacac ggtatcgtta
                                                                              720
    49 ggagcagtag ggacttegge ggtgagtetg ggattaaegg caaattaege aegtaeegga
                                                                              780
    51 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa
                                                                              837
    54 <210> SEQ ID NO: 2
    55 <211> LENGTH: 837
    56 <212> TYPE: DNA
    57 <213> ORGANISM: E. coli
    59 <400> SEQUENCE: 2
    60 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatqtt
                                                                               60
    62 tatgtaaacc ttgcgcctgc cgtgaatgtg gggcaaaacc tggtcgtgga tctttcgacg
                                                                              120
    64 caaatetttt gecataaega ttaeeeggaa aeeattaeag aetatgteae aetgeaaega
                                                                             180
    66 ggttcggctt atggcggcgt gttatctagt ttttccggga tcgtaaaata taatggcagt
                                                                             240
    68 agctatectt tecetactae cagegaaaeg eegegggttg tttataatte gagaaeggat
                                                                              300
    70 aagccgtggc cggtggcgct ttatttgacg cctgtgagca gtgcggggg agtggcgatt
                                                                             360
    72 aaagcagget cattaattge egtgettatt ttgegacaga ccaacaacta taacagegat
                                                                             420
    74 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc
                                                                              480
    76 tgcgatgett etgetegtga tgteaeegtt aetetgeegg aetaeeetgg tteagtgeeg
                                                                             540
    78 attectetta cegittatig igegaaaage caaaaeeigg ggiattaeei eiceggeaea
                                                                             600
```

80 accgcagatg cgggcaactc gattttcacc aataccgcgt cgttttcacc cgcgcagggc

660



DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

82 gtcggcgtac agttgacgcg 84 ggagcagtag ggacttcggc	caacggtacg ggtaagtctg	attattccag	cgaataacac caaattacgc	ggtatcgtta acgtaccgga	720 780
86 gggcaggtga ctgcagggaa	tgtgcaatcg	attattggcg	tgacttttgt	ttatcaa	837
89 <210> SEQ ID NO: 3					
90 <211> LENGTH: 837					
91 <212> TYPE: DNA 92 <213> ORGANISM: E. co	1:				
94 <400> SEQUENCE: 3	11				
95 ttcgcctgta aaaccgccaa	tagtacaget	atccctattq	acaataaasa	agataatat	60
97 tatgtaaacc ttgcgcctgc	cgtgaatgtg	gggcaaaacc	taatcataaa	tetttegaeg	60 120
99 caaatctttt gccataacga	ttatccggaa	accattacag	actatotcac	actocaacoa	180
101 ggctcggctt atggcggcg	t gttatctaai	ttttccqqqa	ccqtaaaata	tagtggcagt	240
103 agctatccat ttccgacca	c cagcgaaac	g ccgcgggtto	r tttataatto	gagaacggat	300
105 aagccgtggc cggtggcgc	t ttatttgaco	g cctgtgagca	qtqcqqqcqq	ggtggcgatt	360
107 aaagctggct cattaattg	c cgtgcttatt	ttgcgacaga	ccaacaacta	taacagcgat	420
109 gatttccagt ttgtgtgga	a tatttacgco	c aataatgatg	r taataatacc	tactggcggc	480
111 tgcgatgttt ctgctcgtg	a tgtcaccgtt	actctgccgg	actaccctqq	ttcagtgcca	540
113 attectetta cegtttatt	g tgcgaaaagc	caaaacctgg	ggtattacct	ctccggcaca	600
115 accgcagatg cgggcaact	c gattttcacc	aataccgcgt	cgttttcacc	agcgcagggc	660
117 gtcggcgttc agttgacgc	g caacggtacg	, attattccca	cgaataacac	ggtatcgtta	720
119 ggagcagtac ggacttcgg	c ggtaagtetg	, ggattaacgg	caaattacgo	acgtaccgga	
121 gggcaggtga ctgcaggga 124 <210> SEQ ID NO: 4	a tgtgcaatcg	attattggcg	tgacttttgt	ttatcaa	837
124 (210) SEQ 1D NO: 4 125 (211) LENGTH: 840					
126 <212> TYPE: DNA					
127 <213> ORGANISM: E. CO	nl i				
129 <400> SEQUENCE: 4	711				
130 ttcgcctgta aaaccgccaa	a tggtaccgca	atccctattq	acaataacaa	caccaatatt	60
132 tatgtaaacc ttgcgcctg	cqtqaatqtq	ggggaaaacc	taatcataaa	tetttegaeg	120
134 caaatctttt gccataacga	ttacccagaa	accattacag	actatotcac	actocaacoa	180
136 ggtgcggctt atggcggcgt	gttatctagt	ttttccggga	ccgtaaaata	taatggcagt	240
138 agctatectt tecetactae	cagcgaaacg	ccacaaatta	tttataattc	gagaacggat	300
140 aagccgtggc cggtggcgct	: ttatttgacg	ccggtgagca	gtgcggggg	agtggcgatt	360
142 aaagctggct cattaattg	cgtgcttatt	ttgcgacaga	ccaacaacta	taacagcgat	420
144 gatttccagt ttgtgtggaa	ı tatttacgcc	aataatgatg	tagtagtacc	cactggcggc	480
146 tgcgatgttt ctgctcgtga	tgtcaccgtt	actctqccqq	actaccctgg	ttcagtgccg	540
148 attectetta cegittatte	rtgcgaaaagc	caaaacctqq	ggtattacct	ctccggcaca	600
150 accgcagatg cgggcaacto	gattttcacc:	aataccgcgt	cqttttcacc	cacacaaaac	660
152 gtcggcgtac agttgacgcg	caacggtacg	attattccag	cgaataacac	ggtatcgtta	720
154 ggagcagtag ggacttcggc	ggtaagtctg	ggattaacgg	caaattacgc	acgtaccgga	780
156 gggcaggtga ctgcagggaa	tgtgcaatcg	attattggcg	tgacttttgt	ttatcaataa	840
159 <210> SEQ ID NO: 5					
160 <211> LENGTH: 840 161 <212> TYPE: DNA					
161 <212> TYPE: DNA 162 <213> ORGANISM: E. CO	1:				
164 <400> SEQUENCE: 5	TI				
	taataaaaa	2++00+5++=			
165 ttcgcctgta aaaccgccaa 167 tatgtaaacc ttgcgcctgc	catastata	accotattg	ycggtggcag	cgctaatgtt	60
169 caaatctttt gccataacga	ttatcccca	yyycadadcc accattacae	Lygicgtaga	cettegacg	120
Junicolde gecaldacya	ccacccyyda	accaltacag	actatgtcac	actgcaacga	180

RAW SEQUENCE LISTING

DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

171 ggctcggctt atggcggcgt gttatctaat ttttccggga ccgtaaaata tagtggcagt	
173 agctatecat the cogactae cagegaact the tector of the cogactae cagegaact cogegaget that a cagegaact the cogactae cagegaget cogactae ca	240
175 aagccgtggc cggtggcgct ttatttgacg cctgtgagca gtgcgggtgg ggtggcgatt	300
177 aaagctggct cattaattgc cattaattat to cetters to the control of the control of the cattaattgc cattaatta to cattaatta the cattaatta to cattaatta the cattaatta to cattaatta the cattaat	360
177 aaagctggct cattaattgc cgtgcttatt 179 gatttccagt ttgtgtgaga tatttccagt ttgtgtgaga tatttccagt	420
179 gatttccagt ttgtgtggaa tatttacget ttgtgacaga ccaacaacta taacagegat 181 tgcgatgtt ctgctcatga tgttgacget aataatgatg tggtggtgcc tactggegge	480
	540
	600
	660
	720
	780
JJJ - JJ J J J J J J J J J J J J J J J	840
	0.10
195 <211> LENGTH: 837	
196 <212> TYPE: DNA	
197 <213> ORGANISM: E. coli	
199 <400> SEQUENCE: 6	
200 ttcgcctgta aaaccgccaa tggcaccgct atccctattg gcggtggcag cgccaatgtt	60
	120
	180
	240
	300
	360
	420
	480
	540
	600
	660
	720
JJJ~~JJ~J~ VL9VQ999QQ LULUCAATCG ATTATTAGGG +~~~~~~~	780
OLY ID NO. /	837
230 <211> LENGTH: 837	
231 <212> TYPE: DNA	
232 <213> ORGANISM: E. coli	
234 <400> SEQUENCE: 7	
235 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatgtt	_
	60
	120
	180
	240
	300
	360
	420
	480
	540
253 attectetta cegittatti tigegaaaage caaaacetig gitattacet eteeggeaca 255 acegeagati eggeaacte gattteaga antanamid	600
255 accgcagatg cgggcaactc gattttcacc aataccgcgt cgttttcacc tgcacagggc 257 gtcggcgtac agttgacgcg caacggtacg attattacca	660
	720
	780
261 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa	837



PATENT APPLICATION: US/09/900,575

DATE: 07/23/2001 TIME: 13:37:27

264 - 212	
264 <210> SEQ ID NO: 8	
265 <211> LENGTH: 837	
266 <212> TYPE: DNA	
267 <213> ORGANISM: E. coli	
269 <400> SEQUENCE: 8	
270 tttgcctgta aaaccgccaa tggcaccgct atccctattg gcggtggcag cgccaatgtt	60
2/2 cacycadact tydedeedd edfaatafa aaacaaaaca faafaafaaa faffaaraa	120
2/4 Cadacetttt geedlaaega ttateeggaa accattacag actatgtoso actagas	180
270 ggoodggct alggcggcgt gttatctaat tittccggga ccgtaaaata tagtagaat	240
270 ayotatooat ttoogactac cagogaaacg cogoggatta tttataatta gagaacaat	300
200 dayccytyyc cydtygcyct ttatttgacg cetgtgagga gtgagggtgg gatara	360
202 adayctyyct Cattaattqc Cqtqcttatt ttqcqacaga ccaacaacta tagaacaa	420
201 gatteccage tigiqiqda tatttacgcc aafaafgafg fggfggfggg tagtggg	480
200 cycyatyttt Ctyctcatda tqtcaccqtt actctqccqq actaccctqq ++qqq+qq	540
200 accepted cognitating the transfer canabide the acceptance of comments	
250 accycladaty cygyclactic gattiticacc aataccgcgt cgttttcacc aggggaggg	600
2)2 yeegyeytac agtigacgcg caacggtacg attattocag cgaataacag ggtatagta	660
234 yyaycaytag ggacttcggc ggtgagtctg ggattaacgg caaattacgg agetaaaaa	720
250 999caygiga cigcagggaa tgtgcaatcg attattggcg tgacttttgt ttata	780
299 (210) SEQ ID NO: 9	837
300 <211> LENGTH: 837	
301 <212> TYPE: DNA	
302 <213> ORGANISM: E. coli	
304 <400> SEQUENCE: 9	
305 ttegeetgta aaacegeeaa tggtacegea atecetatta geggtageag agaanteet	
307 tatgtaaacc ttgcgcctgc cgtgaatgtg gggcaaaacc tggtcgtaga tctttcgacg	60
309 caaatctttt gccataacga ttaccagaa accattacag actatgtcac actgcaacga	120
311 ggttcggctt atggcggcgt gttatctagt ttttccggga ccgtaaaata taatggcagt	180
313 agetateett teeetaetae cagegaaaeg eegegggttg tttataatte gagaaeggat	240
315 aagccgtggc cggtggcgct ttatttgacg ccggtgagca gtgcgggggg agtggcgatt	300
317 aaagctggct cattaattgc cgtgcttatt ttgcgacaga ccaacaacta taacagcgat	360
319 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc	420
321 tgtgatgett etgetegtga tgteacegtt actttgeegg actaecetgg tteagtgeeg	480
323 attectetta cegittatta tycegaaaage caaaacetyg gytattacet atceggeaca	540
325 accgcagate coggcaacte gattttcace aatacegegt cogttttcace egegeaggge	600
327 gteggegtae agttgaegeg caaeggtaeg attatteeag egaataacae ggtategtta	660
329 ggaggagtag ggacttegge ggtaagtata ggatta	720
329 ggagcagtag ggacttegge ggtaagtetg ggattaaegg caaattaege acgtaeegga	780
331 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa 334 <210> SEQ ID NO: 10	837
335 <211> LENGTH: 840	
336 <212> TYPE: DNA	
337 <213> ORGANISM: E. coli	
339 <400> SEQUENCE: 10	
340 ttoggotata accessor to	
340 ttegeetgta aaacegecaa tggcaceget atcectattg geggtggcag egecaatgtt	60
The cary cadaco regedeede cataatata aaaraaaaa taataataa tatta	120
344 CdddCCCCCC gCCdCddCqd TTdCCCCggaa accattacag attatgag actaga	180
340 9900099000 diggoggaattatotaat ttttooggaa oogtaaaata taatu	240
The agreement throughout Carranance concornette the stante and the second	300
350 aageegtgge eggtggeget ttatttgaeg eetgtgagea gtgegggegg ggtggtgatt	360

RAW SEQUENCE LISTING

DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

_ (= : : : : : : : : : : : : : : : : : :	
352 aaagetgget cattaattge egtgettatt ttgegacaga ceaacaacta taacagegat	
	420
	480
	540
	600
	660
	720
	780
	840
370 <211> LENGTH: 837	
371 <212> TYPE: DNA	
372 <213> ORGANISM: E. coli	
374 <400> SEQUENCE: 11	
375 ttcgcctgta aaaccgccaa tggtaccgca ataataa	
377 tatgtaaacc ttgcgcctgc cgtgaatgtg gggcaaaacc tggtcgtaga tctttcgacg	60
	120
	180
	240
385 aagccgtggc cggtggcgct ttatttgacg ccggtgagca gtgcggggg agtggcgatt	300
387 aaagctggct cattaattgc cgtgcttatt there are gtgcgggggg agtggcgatt	360
387 aaagctggct cattaattgc cgtgcttatt ttgcgacaga ccaacaacta taacagcgat 389 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc 391 tgcgatgttt ctgctcgtga tgtcacggtt agtata	420
391 tgcgatgttt ctgctcgtga tgtcacggtt actaltgatg tggtggtgcc cactggcggc	480
391 tgcgatgttt ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgccg 393 attcctctta ccgtttattg tgcgaaaagg gaaaaagg	540
393 attectetta cegittatty tycgaaaage caaaacetyy gytattacet eteeggeaca	600
397 gtcggcgtac agttgacgcg caacggtacg retails	660
397 gteggegtae agttgaegeg caaeggtaeg attatteeag egaataaeae ggtategtta	720
399 ggagcagtag ggacttcggc ggtaagtctg ggattaacgg caaattacgc acgtaccgga	780
401 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa 404 <210> SEQ ID NO: 12	837
405 <211> LENGTH: 840	
406 <212> TYPE: DNA	
407 <213> ORGANISM: E. coli	
409 <400> SEQUENCE: 12	
410 ttcgcctgta aaaccgccaa tggtagaget	
410 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatgtt	60
412 tatgtaaacc ttgcgcccgt cgtgaatgtg gggcaaaacc tggtcgtgga tctttcgacg 414 caaatctttt gccataacga ttatccggaa accattacag actatgtcac actgcaacga 416 ggctcggctt atggcggcgt gttatctaat ttatc	120
416 ggctcggctt atggcggcgt gttatetest the	180
	240
	300
420 aagccgtggc cggtggcgct ttatttgacg cctgtgagca gtgcgggcgg ggtggcgatt	360
	420
424 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc tactggcggc	480
426 tgcgatgttt ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgcca 428 attcctctta ccgtttattg tgcgaaaagg actaccctgg ttcagtgcca	540
428 attectetta cegittatty tycgaaaage caaaacetyy gytattacet eteeggeaca 430 acegeagaty cyggeacte gatttteses arts	600
	660
432 gtcggcgtac agttgacgcg caacggtacg attattccag cgaataacac ggtatcgtta 434 ggagcagtag ggacttcggc ggtgagtatc	720
434 ggagcagtag ggacttegge ggtgagtetg ggattaacgg caaattatge acgtacegga 436 gggcaggtga ctgcagggaa tgtgcaateg attatte	780
436 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaataa 439 <210> SEQ ID NO: 13	840
440 <211> LENGTH: 837	
441 <212> TYPE: DNA	
III. Dun	

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/900,575

DATE: 07/23/2001
TIME: 13:37:28

Input Set : A:\FimH Peptides.ST25.txt
Output Set: N:\CRF3\07232001\1900575.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date L:2195 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:46 L:2453 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:47 L:2489 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:48 L:2531 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:49 L:2569 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:50 L:2654 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:52 L:2732 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:55 L:2796 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:56 L:2808 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:57 L:2820 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:58 L:2832 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:59 L:2844 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:60 L:2856 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:61 L:2868 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:62 L:2880 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:63 L:2892 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:64